

Rocky Intertidal Monitoring Protocol for the Redwood National and State Park, CA

Standard Operating Procedure (SOP) # 13: Metadata Guidelines

Version 1.00 (March 2008)

Revision History Log:

Previous Version	Revision Date	Author	Changes Made	Reason for Change	New Version
			-		

This Standard Operating Procedure (SOP) provides information for adding metadata documentation from the Redwood National and State Parks (RNSP) Rocky Intertidal Monitoring program. The MARINe database serves both as a key record and as a digital repository for all intertidal monitoring-related data, while metadata preserves valuable information (the “who, what, where, when, and how”) associated with each record in a concise format. Metadata can be thought of as “data about data,” a set of information that describes (in full) each data file (e.g., photograph) in the archive. This SOP is adapted from the Klamath Network (KLMN) Metadata Guidelines, written by KLMN Data Manager, Sean Mohren (Mohren 2007a). The guidelines are written to reflect current best practices used by the Klamath Inventory and Monitoring (I&M) Network, with standards developed for implementation (Mohren 2007b). While software and network applications will evolve, the intent and content of metadata should remain relatively consistent.

I. General Considerations

Metadata is meant, among other uses, to describe the content, quality, condition, and other characteristics of data files. Clear, complete, descriptive, and standardized metadata is important, as this improves the efficiency of finding data of interest. Cataloging accurate and complete metadata is a powerful technique to preserving critical information necessary for secondary data analysis and for cooperative data and image sharing.

Metadata collection is mandated for federal agencies by Executive Order 12906, signed in June 1994. Efforts to standardize metadata usage across NPS is expanding.

II. Introduction to Metadata

Data documentation is a critical step towards ensuring that datasets are usable for their intended purposes well into the future. This involves the development of metadata, which is defined as structured information about the content, quality, condition, and other characteristics of a dataset.

SOP #13: Metadata Guidelines (continued).

In addition to spatial information, metadata includes information about data format, collection and analysis methods, time of collection, originator, access/use constraints, and distribution. Metadata provides the means to catalog datasets within Intranet and Internet systems, making the associated datasets available to a broad range of potential users. While most frequently developed for geospatial data, metadata describing non-geospatial datasets is also needed (NCCN 2006a).

III. Timelines

It is the responsibility of the Project Manager, working with the Data Manager, to define timelines for metadata delivery during the planning phase of a project. If no timelines are defined, then it is the Project Manager's duty to submit metadata or metadata products (e.g., Metadata Interview Form) to the Klamath Network Data Manager in the proper format, no less than one month prior to the start of a new field season.

IV. Responsibilities and Standards

Metadata is an important piece of documentation that helps guarantee the long-term usability of data. The degree of documentation will vary depending on the product, but a few standards will always hold true.

1. Data collected through I&M funded projects will meet FGDC, NBII, and National Park Service (NPS) standards before being made available to the public.
2. Project Managers will be expected to submit a data dictionary and Metadata Interview Form prior to the start of the first field season.
3. Project Managers will be expected to review and revise the data dictionary and Metadata Interview Form at the end of each field season and report changes following the timeline listed in SOP #17: Project Deliverables.
4. It is the responsibility of the Data Manager to develop the official metadata based on the data dictionary and Metadata Interview Form provided by the Project Manager (Table 1).
5. It is the Data Manager's responsibility to parse and transfer metadata to the NPS Data Store, if applicable.
6. The Data Manager will work with the Project Manager and park staff to determine the sensitivity level of any data.

SOP #13: Metadata Guidelines (continued).

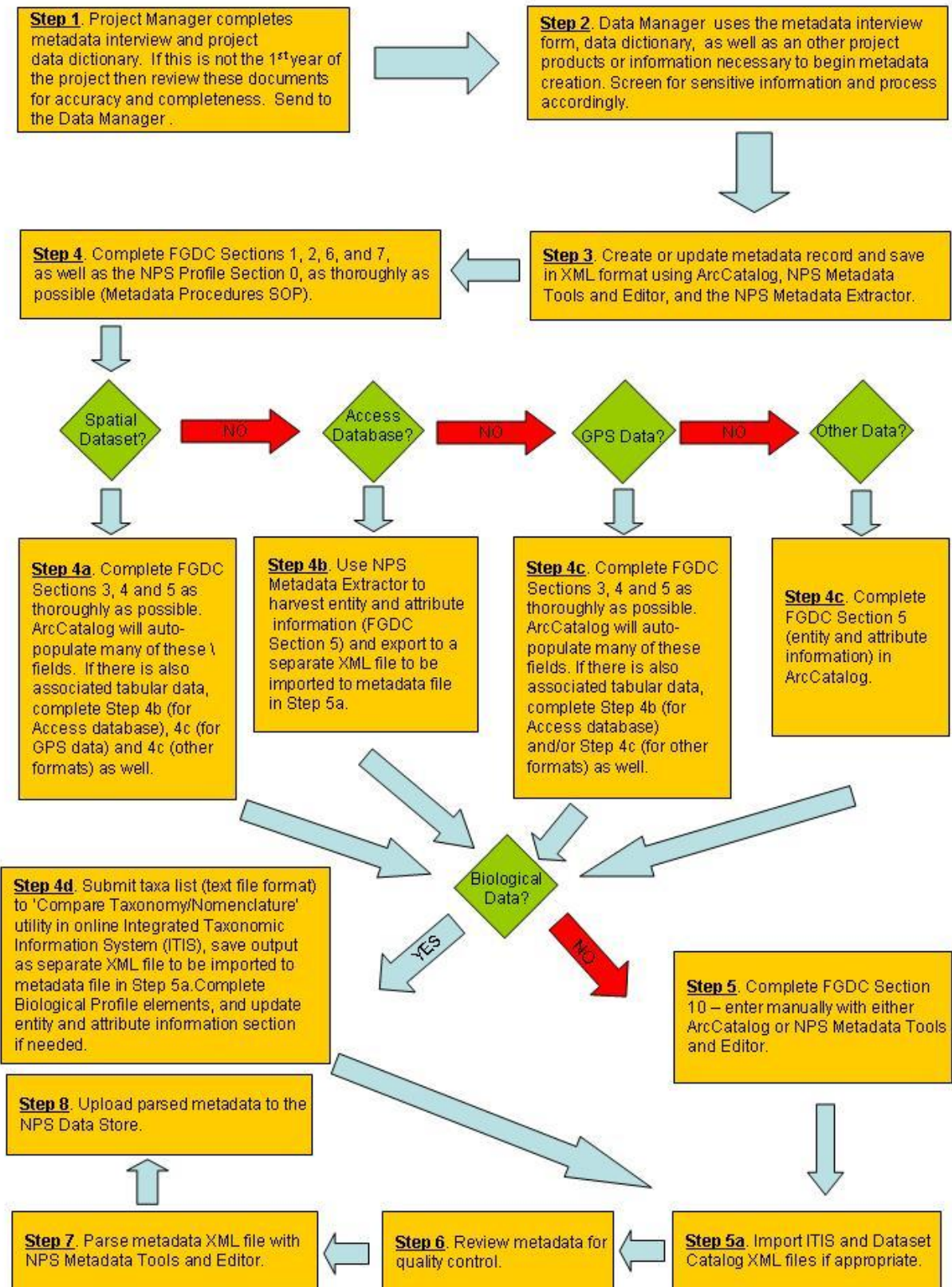


Figure 1. The following diagram shows the general workflow for metadata creation.

SOP #13: Metadata Guidelines (continued).

V. Steps for Metadata Creation

Step 1: Metadata Interview Form and Information Gathering

- A. The Project Manager should obtain and complete the KLMN Metadata Interview Form at project onset. This facilitates compiling the information required to create compliant metadata.
 - 1. The KLMN Metadata Interview Form is posted on the KLMN Internet and Intranet web pages. In addition, the form can be obtained by contacting the Network Data Manager.
 - 2. Best attempts should be made to populate the Metadata Interview Form as completely as possible prior to starting field work. However, it is recognized that changes to the form will occur throughout the project.
- B. A data dictionary must be created to provide information to help the Data Manager create or update the official metadata and, in some cases, the project database. The data dictionary should include:
 - 1. The name and purpose of each table, shapefile, coverage, or feature class.
 - 2. A list of attribute names, type, size, and description by table.
 - 3. If the database was not designed by the Klamath Network, it is the responsibility of the Project Manager to provide the Data Manager with the data dictionary.
- C. If a taxa list other than the current ITIS certified taxa list was used, the Project Manager will need to provide that list at the end of each field season. Taxa lists should include:
 - 1. Taxon group (Bird, Mammal, Reptile, Invertebrate, etc.).
 - 2. Scientific name.
 - 3. Common name.
 - 4. Any special code that defines a species.
- D. The Project Manager should send a copy of any additional information that might be valuable for the development of metadata.

Step 2: Sensitivity Review

Sensitive data (species locations, site locations, etc) may not be subject for release to the public.

- A. The current version of the NPS Data Store does not screen for sensitive information. Therefore, any data with a sensitive status will not be posted on the Data Store.
- B. The Network Data Manager will be responsible for posting data as sensitive. Status of the data will be based on comments provided by the Project Manager under the “Sensitivity” question on the Metadata Interview Form. In addition, the KLMN will consult with park staff if the sensitivity status of any data is questionable.

Step 3: Metadata Software Selection

- A. The Klamath Network will utilize ArcCatalog, NPS Metadata Tools and Editor, and the Database Metadata Extractor to create metadata for all projects.
 - 1. ArcCatalog automatically harvests spatial organization and reference information, as well as entity and attribute information for GIS datasets.
 - 2. The NPS Metadata Tools and Editor is provided as a stand alone program or as an extension for ArcCatalog and is available at:

SOP #13: Metadata Guidelines (continued).

<http://science.nature.nps.gov/nrdata/tools/>

- i. It can be used for metadata creation and editing.
 - ii. You can use this tool to import, export, and parse metadata
 - iii. You cannot harvest entity and attribute information. However, this is an anticipated feature for the next version.
3. The [NPS Database Metadata Extractor](#) is a custom software application for authoring, editing, and managing NPS metadata. The metadata extractor operates either as an extension to ArcCatalog versions 8.3/9.x or as a stand alone desktop application. Eventually this tool will be incorporated into the NPS Metadata Tools and Editor. Features of this tool include:
 - i. Automatically harvests entity (table) and attribute (field) metadata from MS Access databases, including domains.
 - ii. Allows the user to edit and review the harvested metadata and make batch edits.
 - iii. Allows the user to export metadata to a FGDC-compliant XML file.
 - iv. Exported XML can be used in the Metadata Tools and Editor either by opening it to start a new metadata record or by updating it with the template to fill in FGDC Section 5 of an existing metadata record.

Step 4: Additional Requirements

- A. Along with the required metadata, the Klamath Network requires the following information be included in the metadata document:
 1. The name and agreement code for the project. These references can be entered in the Related Key element in the Program Information Section (NPS Section 0) on the NPS Profile.
 2. References to all products (GIS, GPS, Databases, Reports) generated by the projects. These references can be entered in the repeatable Cross Reference element of the Identification Information Section.
 3. Standard NPS liability language in the Distribution Liability metadata element of the Distribution Information (FGDC Section 6).
 - i. This can be found at: <http://www.nps.gov/gis/liability.htm>.

Step 5: Biological Data Profile

If a dataset includes biological information, the Biological Data Profile provides a set of extended metadata elements to document the species observed, taxonomic information, methods, and analytical tools.

- A. The most direct, and KLMN preferred, means to populating the Biological Data Profile metadata elements are outlined in [Biological Profile \(National Biological Information Infrastructure - NBII\) Metadata Guide](#) (NPS Data Store 2005a).
 - i. This approach primarily utilizes the NPS Metadata Tools and Editor and may also require the entity and attribute harvesting capability of NPS Database Metadata Extractor for Access datasets.
- B. The following two guidance documents describe alternative approaches to completing the Biological Data Profile for a metadata record. Note that the first requires the use of additional metadata creation software (Spatial Metadata Management System, or SMMS):

SOP #13: Metadata Guidelines (continued).

- i. [Metadata Tools Used in the Creation of the FGDC Biological Data Profile](#) (Callahan and Devine 2004).
- ii. [National Biological Information Infrastructure \(NBII\) Metadata Steps](#) (McGuire 2004).

Step 6: Metadata Review

Review metadata for quality control (QC) prior to posting to NPS Data Store. A useful QC Checklist is available for download on the NPS Intermountain Region GIS website:

http://imgis.nps.gov/tips_templates.html.

Step 7: Metadata Parsing and Exporting to XML Format

The NPS Data Store requires that metadata records be parsed into FGDC-structured metadata and then exported to XML format.

- A. If using ArcCatalog, these steps can both be done directly with the NPS Metadata Tools and Editor. See [Parsing Metadata with the NPS Metadata Tools and Editor](#) (NPS Data Store 2005b) for more information.
- B. If using another application, export the metadata first to ASCII text format and then parse with Metadata Parser (MP). MP can simultaneously output an XML format metadata file as well.
 - i. MP must be customized to handle NPS, Biological Data, or ESRI Profile metadata elements. For specifics, refer to:
 - a. The README.txt file included in the zipped NPS Metadata Profile configuration files available from the NR-GIS Metadata and Data Store website at:
<http://science.nature.nps.gov/nrdata/docs/metahelp/metahelp.cfm>.
 - b. [Parsing Metadata with the NPS Metadata Tools and Editor](#) (NPS Data Store 2005b)

Step 8: Metadata Posting

Post the metadata to the NPS Data Store.

- A. Authorized NPS staff may request upload and edit access to the NPS Data Store through the NPS Natural Resource Universal Web Login (UWL) available at: <https://science1.nature.nps.gov/nrdata/>. This is also the portal for uploading data.
- B. More information about metadata upload format requirements is available at: <http://science.nature.nps.gov/nrdata/docs/metahelp/metainfo.cfm> and in [Metadata and Data Uploading Guidance](#) (NPS Data Store 2005c).

Step 9: Editing/Updating Metadata Already Posted to NPS Data Store

As of Version 1, the NPS Data Store application allows online editing of NPS Theme Category and ISO Theme Keyword information and the deletion of single metadata records and/or datasets only (help documentation: <http://science.nature.nps.gov/nrdata/docs/metahelp/edithelp.cfm>).

- A. For metadata records simply needing edits to NPS Theme Category or ISO Theme Keyword elements, refer to [Editing Category Information](#) (NPS Data Store 2005d).
- B. If a metadata records posted to the NPS Data Store contain errors or require edits to other elements, you should contact the Klamath Network Data Manager. Erroneous metadata will need to be deleted from the NPS Data Store, edited, then reposted. Refer to [Deleting Single Records](#) (NPS Data Store 2005e).

SOP #13: Metadata Guidelines (continued).

- i. The user should first download the metadata record (save in XML format) to the local system, then edit as needed in a text editor or metadata software program.
- ii. The edited metadata record can then be resubmitted to the NPS Data Store.
- iii. If the dataset documented by the metadata record requires no edits, it will not need to be reposted. Simply ascertain that the metadata file still specifies the correct pathway to the dataset on the NPS Data Store before resubmitting the metadata file.

VI. Literature Cited

- Callahan, K., and H. Devine. 2004 (draft). Metadata tools used in the creation of the FGDC biological data profile. National Park Service, Northeast Region. Online. (http://science.nature.nps.gov/im/datamgmt/docs/SOP_BioDataProfileTools_v1.doc). Accessed 30 January 2008.
- Executive Order No. 12906 (Coordinating geographic data acquisition and access: The national spatial data infrastructure). April 11, 1995, 59 FR 17671.
- McGuire, S. 2004. National Biological Information Infrastructure (NBII) metadata steps. National Park Service, Midwest Region. Online. (<http://www.nature.nps.gov/im/units/mwr/gis/download/BiologicalProfileGuide.pdf>). Accessed 30 January 2008.
- Mohren, S. R. 2007a. Data management plan, Klamath Inventory and Monitoring Network. Natural Resource Report NPS/KLMN/NRR--2007/012. National Park Service, Fort Collins, CO.
- Mohren, S.R. 2007b. Metadata guidelines. Klamath Inventory and Monitoring Network, National Park Service
- North Coast and Cascades Network – National Park Service. 2006b. Metadata procedures SOP. National Park Service. Online. (http://www1.nature.nps.gov/im/units/nccn/dm_docs.htm). Accessed 6 November 2007.
- North Coast and Cascades Network – National Park Service. 2006c. NCCN metadata interview. National Park Service. Online. (http://www1.nature.nps.gov/im/units/nccn/dm_docs.htm). Accessed 6 November 2007.
- NR-GIS Data Store. 2005a. Biological profile (National Biological Information Infrastructure - NBII) metadata guide. National Park Service, Natural Resource and GIS Programs. Online. (<http://science.nature.nps.gov/nrdata/docs/metahelp/BiologicalProfileGuide.pdf>). Accessed 30 January 2008.
- NR-GIS Data Store. 2005b. Parsing metadata with the NPS Metadata Tools and Editor. National Park Service, Natural Resource and GIS Programs. Online. (<http://science.nature.nps.gov/nrdata/docs/metahelp/NR-GISMetadataParsingGuidance.pdf>). Accessed 30 January 2008.

SOP #13: Metadata Guidelines (continued).

NR-GIS Data Store. 2005c. Metadata and data uploading guidance. National Park Service, Natural Resource and GIS Programs. Online. (<http://science.nature.nps.gov/nrdata/docs/metahelp/NR-GISMetadataDataUploadGuidance.pdf>). Accessed 30 January 2008.

NR-GIS Data Store. 2005d. Editing category information. National Park Service, Natural Resource and GIS Programs. Online. (<http://science.nature.nps.gov/nrdata/docs/metahelp/NR-GISDataStoreEditCategory.pdf>). Accessed 30 January 2008.

NR-GIS Data Store. 2005e. Deleting single records. National Park Service, Natural Resource and GIS Programs. Online. (<http://science.nature.nps.gov/nrdata/docs/metahelp/NR-GISDataStoreDeleting.pdf>). Accessed 30 January 2008.

SOP #13: Metadata Guidelines (continued).

Electronic Metadata Interview

[Note: Please make your responses directly within this word document in "Red" type.]

1. Have you already prepared metadata for this data set?
 - a. If yes, please send a copy of the document or reference to where it can be found and skip to item **18**.
2. What is the title of the data set?
3. Who are the originator(s)/owner of the data set? (Include address, and telephone number)
 - a. If someone else should answer question about the data, please list the name, address, and telephone number.
 - b. Are there other organizations or individuals who should get credit for support, funding, or data collection and analysis?
4. Does the data set contain any sensitive information that should not be released to the public?
NPS?
 - a. Explain why the data should not be released to the public.
 - b. Explain why the data should not be released to non-park NPS staff.
5. Is the data set published or part of a larger publication?
 - a. If so, what is the reference?
6. Include a brief (no more than a few sentences) description of the data set.
7. Why were the data collected in the first place?
8. What is the time period represented by the data set?
9. Where the data developed primarily through:
 - a. Field visits.
 - b. Remote instrumentation (i.e., Temperature recorders, etc.).
 - c. Existing data sources.
10. What is the status of the data you are documenting? – *complete, in progress, planned*
 - a. Will the data set be updated? If so, how frequently?
11. Where were the data collected? Include description and coordinates, if known.
12. List some keywords to help search for this data set.
 - a. Thematic, Place, Temporal, Strata, Taxonomy.
 - b. If a controlled vocabulary was used, what is the reference?
13. List any related data sets that could be documented for cross-reference.
14. The FGDC Biological Profile included the means to document tabular data sets, taxonomy, field methods, and the use of analytical tools or models.
 - a. Was your data set developed using a model or other analytical tool?

SOP #13: Metadata Guidelines (continued).

- i. If so, what is the reference?
 - ii. If the model or tool is available include a contact and/or URL.
 - b. Does the data set contain biological information? If no, skip to item **15**.
 - i. What species or communities were examined?
 - ii. Did you use a taxonomic authority or field guide for identification? If so, what is the reference?
 - iii. Briefly summarize your field methods (cut & paste from other documents!).
 - 1. If you used existing protocols or methods, list the references.
 - iii. If you use a different taxonomic hierarchy than what is available in ITIS then you need to supply the taxonomic hierarchy for all species within the data set.
- 15. Is your data set archived in a databank or data catalog? If yes, please include a reference to the documentation and skip to item 16. If No:
 - a. What Measures did you take to make certain that your data set was as nearly correct as possible?
 - b. Were there any things that you excluded from your data collection i.e., stems less than a certain diameter or streams without surface flow.
 - c. What is the form of your data set? - *spreadsheet, ASCII file, gis layer, database, other*.
 - d. What is the filename for your data set?
 - i. For each file or table, list the fields in the data set and for each field list:
 - ii. The definition of the field
 - iii. If the data are coded (Enumerated Domain), list the codes and the definitions.
 - iv. If the codes come from a published code set (Codeset Domain), list the reference.
 - v. If the data are measured (Range Domain), list the units and the minimum and maximum allowable values ("no limit" is acceptable).
 - vi. Otherwise, the domain is unrepresentable. Include a brief description of what is in the field.
- 16. Is this a GIS data set? If no, skip to item 17.
 - a. Include a path to where the data can be accessed over the network or send a copy of the ArcInfo export file, an ArcView shapefile, or an ArcCatalog exported metadata file (txt or xml).
 - i. Include projection parameters, if necessary.
 - b. List any source data sets you used. For each source list:
 - i. Source name, originator and publication date.
 - ii. Source time period and scale.

SOP #13: Metadata Guidelines (continued).

- iii. Source presentation form and media type.
 - iv. Contribution of source to your analysis.
 - c. List the processing steps you used to create your data set, including the approximate date of processing.
17. Is the data set available for distribution? If no, go to 18.
- a. Are there legal restrictions on who may use the data?
 - b. Do you have any advice for potential users of the data set?
 - c. What are your distribution instructions?
18. You are done. Send this completed document with the relevant responses to this interview to your metadata coordinator (Sean Mohren, Klamath Network Data Manager. Sean.Mohren@nps.gov, 541-552-8576).